Are More Important Patents Approved More Slowly and Should They Be?*

Pierre Régibeau and Katharine Rockett

University of Essex

June 2003

Preliminary

Abstract

Innovative activities often are heavily regulated. Reviews conducted by administrative agencies take time and are not perfectly accurate. Of particular concern is whether, by design or not, such agencies discriminate against more important innovations by taking more time to perform their reviews. We study the relationship between the length of patent review and the importance of inventions in a theoretical model. We find that, controlling for the importance of innovations, the welfare-maximising patent approval delay decreases over time. Second, controlling for a patent's position in the new technology cycle, the optimal examination time decreases with the importance of patents. We test our predictions on US GM crop patent data from 1988 to 1998. The evidence supports the predictions of the theoretical model.

*The authors would like to acknowledge the generous support of ESRC grant L145251003 for this project.

Innovative activities are often heavily regulated. In some cases regulation affects the conditions under which research is conducted: most industrialised countries have strict rules about lab safety and how human or animal subjects taking part in experiments ought to be treated. In other cases, regulation determines the conditions under which a new product can actually be introduced into the market. A prominent example is the U.S. Food and Drug Administration that decides whether and for what uses new prescription drugs can be commercialised. More recently, genetically modified (GM) crops have also been subjected to complex trial and approval procedures. In fact, the commercial exploitation of such crops has so far only been approved in Argentina, Canada and the U.S.[1]

Less explicit but equally important is the 'regulation' of innovation through the allocation of intellectual property rights. In the patent approval process, for example, professional examiners from the patent office evaluate patent applications thoroughly in order to determine whether the innovations that they claim meet the usual criteria of novelty, non-obviousness and usefulness.

Reviews conducted by agencies like the FDA and the USPTO are administrative processes that take time and cannot claim to be perfectly accurate. In fact, the accuracy and speed with which such agencies perform their task can have significant effects on the amount and type of research undertaken by economic agents. Of particular concern is whether, by design or not, these institutions discriminate against more important innovations. This question was first asked by Dranove and Metzler (1994) who examined drugs approval by the US FDA. They found that important drugs actually make it to market significantly faster than less important ones. On the contrary, Johnson and Popp (2003) found that, over the period 1976 – 1996, important US patents have on average been processed more slowly than less important ones. They went on to argue that, as US patent applications have recently become available 18 months after their filing date, the extra approval delay experienced by important patents could erode their relative profitability biasing private research incentives. Moreover, given the patent examination bias, the recent

[1] At the end of May 2003, the Indian government also announced that it would approve the cultivation of protein-enhanced GM potatoes for distribution to children of poor families.

introduction of early disclosure of patent information could be socially detrimental unless the benefits from faster diffusion of knowledge are large.

In this paper, we revisit the relationship between the length of patent review and the importance of inventions using US patent data on GM crop innovation. We first specify a simple model of the patent approval process. A key feature of this model is that patent granting decisions are imperfect but their precision can be improved by more thorough examination of the applications. Hence longer approval delays make for better decisions. Good decisions are important for two reasons. Firstly mistakenly granting patent rights amounts to granting some monopoly power without the benefits of technological improvements or diffusion. Another important aspect of the model is that technological uncertainty decreases over time. This is supposed to capture the idea that industries are characterised by innovation cycles. A new cycle begins when fundamentally new technological routes are explored. As the cycle unfolds, patent examiners become more familiar with the new approaches and the precision of patent decisions improves.

We obtain two main results. Firstly, controlling for the importance of innovations, the welfare-maximising patent approval delay decreases over time. Secondly, controlling for a patent 's position in the new technology cycle, the optimal examination time decreases with the importance of patents. In spite of these clear-cut results we show that testing the relationship between delay and patent importance on a cross section of fields of innovation could easily produce a *positive* correlation between the two variables. This is due the combination of two factors. Firstly, at a given time, different fields of innovation are bound to be at different point of their 'cycles'. Secondly, major innovations tend to emerge disproportionally near the beginning of such cycles. Unless one can control precisely for the patents 'respective fields of innovation then, one will tend to find more important innovations in areas that are in the early phases of a new technology cycle…which is also when patent approval delay should – optimally – be longer.

To address this issue we use a data set containing most of the US patents relative to GM crops from 1988 to 1998. This is a well-defined research field and the sample period covers the first ten years of a new technological cycle. We can therefore hope to disentangle the effects of both 'time' and 'importance' on the examination delays experienced by patents. Although the empirical part of the paper is best viewed as a rough first look at the data, the evidence so far supports the predictions of our model. Controlling for the importance of patents, the time required to obtain approval decreases significantly over time. Although the results are less clear-cut, there is also some support for the claim that, controlling for the time of patent application importance and delay are negatively correlated.

The rest of the paper is organised as follows. We describe our model in section I and characterise the welfare-maximising examination delay in section II. Section III outlines the empirical implications of our model and briefly present the data. Our empirical results are explained in section IV and section V concludes.

I. A simple model of optimal patent approval delay

Potential innovators can invest in two types of projects, x and y. Type x projects do not produce inventions that meet the traditional patenting criteria of novelty, non-obviousness and usefulness while type y projects do. However, the patenting process is imperfect. For each 'invention' reviewed, there is a probability P that the examiner will reach the wrong conclusion, namely grant a patent to an x project or not grant it to a y project.[2]

A y project involves an initial outlay of y. If the patent is granted, the innovator gets private benefits $\pi(y)$. These are total benefits discounted to the time at which the patent is granted. If the patent is not granted, the innovators' private benefits are equal to 0. The total discounted social welfare (non inclusive of the initial outlay y) created by an innovation y is W(y) if a patent is granted and W(y,0) if a patent is not granted. We assume that:

---

[2] We assume that the same probability of error applies to both types of projects to keep computations simple. Extending the analysis to the case where the probability of error differs across project types is messy but straightforward. It does not convey any additional economic insights.

$$W(y) > \pi(y) > y$$

$$W(y,0) > W(y) > 0$$

The first inequality means that the innovator fails to appropriate the whole surplus generated by its invention. Essentially, this amounts to assuming that the 'business stealing' effects of the invention are not too large. We also assume that y innovations are both socially and privately worthwhile. The second inequality simply reflects the fact that, once the innovation has been achieved, society actually prefers not to grant protection to avoid the corresponding monopoly distortion.[3]

An x project does not involve any initial outlay. If no patent is granted than private and social benefits are both equal to zero, i.e.:

$$\pi(x,0) = W(x,0) = 0$$

If a patent is in fact – mistakenly – obtained then the 'inventor' enjoys positive private benefits but social benefits are negative:

$$W(x) < 0 < \pi(x) < \pi(y)$$

One possible interpretation of this assumption is that the positive private benefits do not come from using a valuable innovation – there is not any – but from the ability to exclude others from a given area of research.[4] $W(x) < 0$ just reflects the social cost of such gratuitous exclusionary behaviour. While we find this assumption appealing, it is not strictly necessary. Our basic conclusions would be unchanged as

[3] This assumption will only hold if the social benefits of revealing information through patent documents are not too large. If they were, then society might be better off with a period of strict monopoly but information diffusion than with a period of lesser monopoly (e.g. when the patent is exploited as a less easily enforceable trade secret) without diffusion of information. Anyway, as will become clear later, this assumption is only used to keep the analysis as simple as possible. Reversing the inequality would not significantly affect our conclusions.

[4] These can be benefits from preventing rivals from obtaining their own (legitimate) property rights or the x innovator might be able to obtain royalties from others whose genuine innovation might infringe on the x innovator 's ill-obtained property rights.

long as Net welfare is higher for a project y than for a project x, i.e. as long as W(y) – y > W(x).

We now turn to the patent-granting process. Examiners cannot readily identify whether a patent application refers to a legitimate invention y or to a 'non-invention' x. Determining the type of the application is precisely the object of the extensive reviews to which patent applications are subjected. Such reviews take time. The more time it takes, the more thorough the review and, therefore, the lower the probability of error. Using our notation, the probability of error can be written as

$$P = P(d,t)$$

, where P is continuously decreasing in d, the time that elapses between the date of filing and the date at which a patent is granted or denied. We will make the conventional assumption that the marginal benefit of waiting is decreasing and that error cannot be completely eliminated in finite time, i.e.:

$$\frac{\partial^2 P}{\partial d^2} > 0 \ \text{ and } \lim_{d \to \infty} P(t,d) = 0$$

Finally we will assume that the inititial learning from inspection is large enough, i.e.:

$$-\frac{\partial P}{\partial d}(d = 0) > \frac{r}{2}$$

,where r is the discount rate. The probability of erroneous judgement, P, is also a continuous function of a time variable t. This variable measures the time elapsed since the beginning of a new innovation 'cycle'. A new cycle begins when research takes a radically new direction. One can think, for example of the discovery of an entirely new class of therapeutic drugs, the first genetic manipulations of crop plants or the first PC-based application software. Radical change often means significant uncertainty as to the nature and validity of the inventions claims submitted to the Patent Office. Such uncertainty is likely to decrease over time as a greater number of applications provide better points of reference and as patent examiners become more familiar with the science involved. This type of learning is introduced in the simple possible way by assuming that the probability P of erroneous patent decisions is decreasing in t.

The profit from investing in an x-type project, discounted at time 0 is:

$$\Pi(x) = e^{-rd}P(d,t)\pi(x)$$

, where r is the discount rate. This simply says that x projects only make profits when the examiner mistakenly grants a patent. The discounted profit from investing in a y-type project is:

$$\Pi(y) = -y + e^{-rd}\,[1 - P(d,t)]\,\pi(y)$$

reflecting the fact that a patent is granted if the examiner does not make a mistake. Therefore, private investors will prefer to pursue legitimate innovation projects if and only if:

$$y \le e^{-rd}\,[\pi(y) - P(d,t)(\pi(y) + \pi(x))]$$

In order to illustrate this decision rule graphically, we must first determine the precise shape of the $\Pi(x)$ and $\Pi(y)$ functions. We have:

$$\frac{\partial \Pi(x)}{\partial d} = \pi(x)e^{-rd}[-rP(d,t) + \frac{\partial P}{\partial d}] < 0$$

$$\frac{\partial^2 \Pi(x)}{\partial d^2} = \pi(x)e^{-rd}[r^2 P(d,t) - 2r\frac{\partial P}{\partial d} + \frac{\partial^2 P}{\partial d^2}] > 0$$

i.e. $\Pi(x)$ is a continuously decreasing and convex function of d.

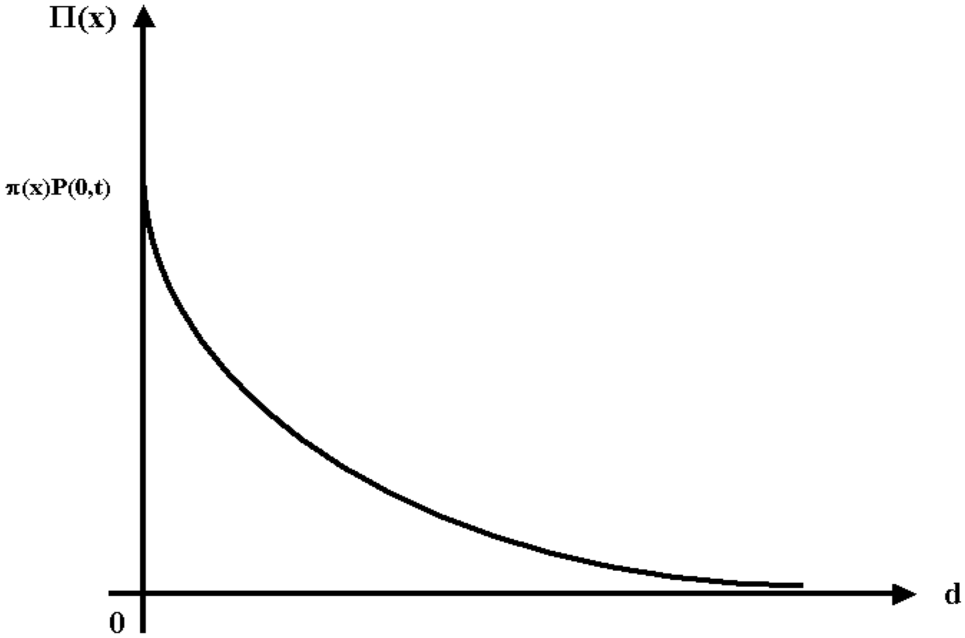

Figure 1

Similarly, we have:

$$\frac{\partial \Pi(y)}{\partial d} = \pi(y)e^{-rd}[-r(1-P(d,t)) - \frac{\partial P}{\partial d}]$$

The first term inside the bracket is negative, reflecting the fact that a longer time of approval reduces the discounted values of the expected profits. The second term inside the bracket is positive: a longer delay means greater precision in the examiner 's decision which, for a y investor, is good news. Moreover, given our previous assumption on the minimum magnitude of $\frac{\partial P}{\partial d}(d=0)$, we have that

$$\frac{\partial \Pi(y)}{\partial d}(d=0) > 0$$

As we also have $\lim_{d\to\infty} \Pi(y) = -y$, we know that Π(y) initially increases in d and ultimately approaches –y asymptotically from above. Although not essential to the argument, conditions for Π(y) to have a single maximand d_{max} are given in the appendix. Assuming that these conditions are satisfied Π(y) can be represented as follows:

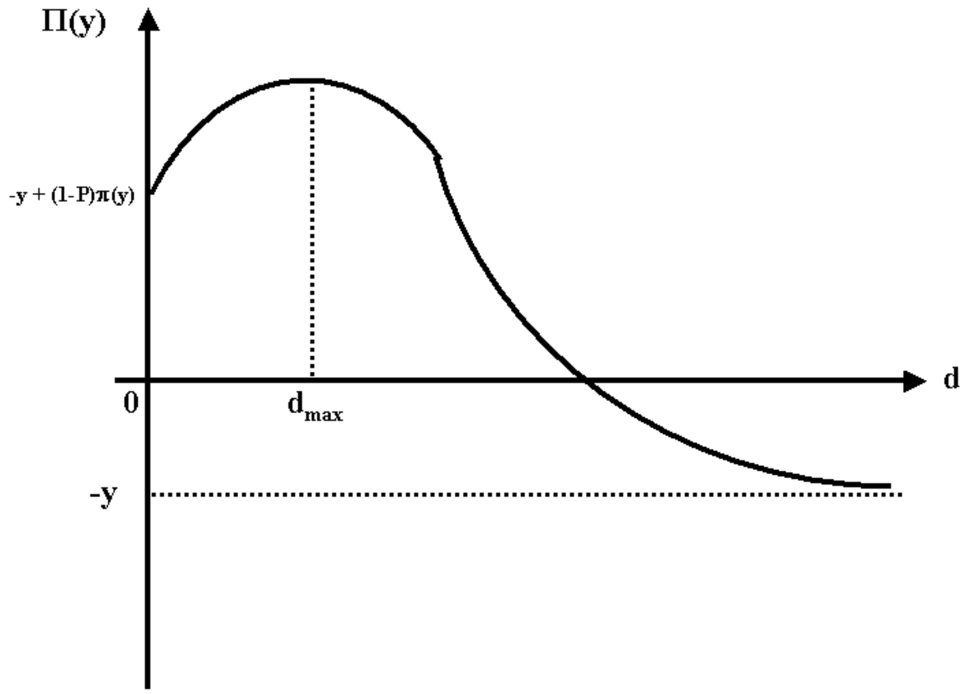

Figure 2

II. Welfare-Maximising Delay

Comparing figures 1 and 2 reveals the importance of the relative size of the vertical intercepts of the two curves. We will ignore the case where Π(x) lies everywhere above Π(y) as investors never engage in worthwhile innovation activities.

Consider the case where the probability of error at d = 0 is low enough to ensure that, absent any examination delay, y would be more profitable than x (Figure 3.a.). This ensures that the two curves will intersect (at least) once. However, these points of intersections are quite irrelevant to establish the socially optimal time of approval. Since the socially worthwhile project will be chosen anyway, there is no need for a positive delay. Increasing the delay beyond d = 0 would only reduce welfare, in two ways. Firstly, it would unnecessarily push back the date at which the social benefits of the innovation can be reaped. Secondly, it would increase the precision of the examiner ' s decision, increasing the chance that a patent is granted. This is undesirable since, conditional on y being chosen, welfare is higher in the absence of patent protection.

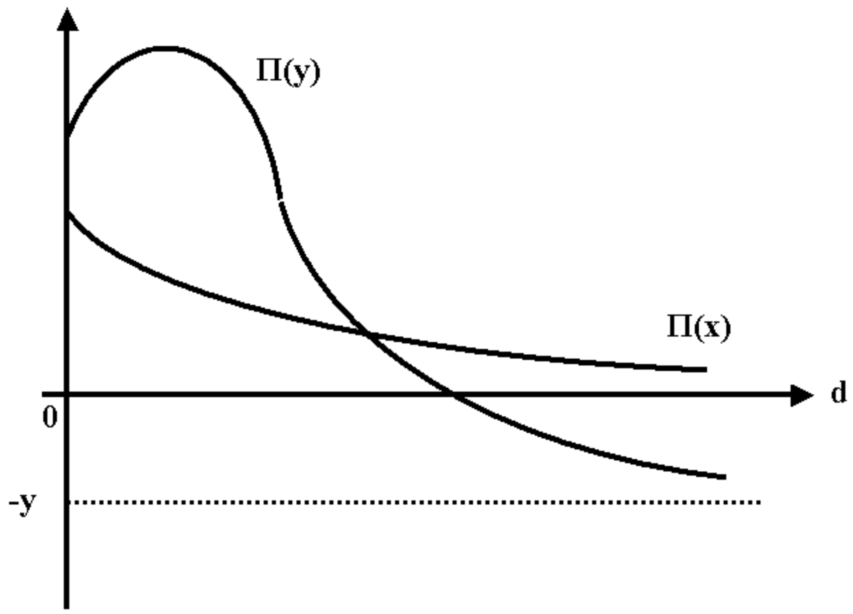

Figure 3.a.

Now, consider the case where the probability of error at d = 0 is sufficiently high to ensure that, absent any examination delay, x would be more profitable than y, but not high enough to prevent the two curves from intersecting (Figure 3.b.). This time, there must be at least two points of intersection. However, following the same argument as above, only the leftmost intersection would be relevant to the welfare maximisation problem anyway. Hence, the socially optimal delay that triggers investment into y is d*. Moreover, as x actually decreases welfare, triggering investment into y is desirable. Hence d* is actually the welfare-maximising delay.

Proposition 1: There exists an intermediate ranges of values of P(0,t) for which the socially optimal time of approval is strictly positive.

Ultimately, what we are trying to determine is whether the positive correlation between patent "importance" and time of approval reported in Johnson and Popp (2003) can be compatible with a socially optimal patent review policy. In order to do this we need to know how the welfare maximising delay, d*, changes over time and with the "importance" of the socially worthwhile innovation y.

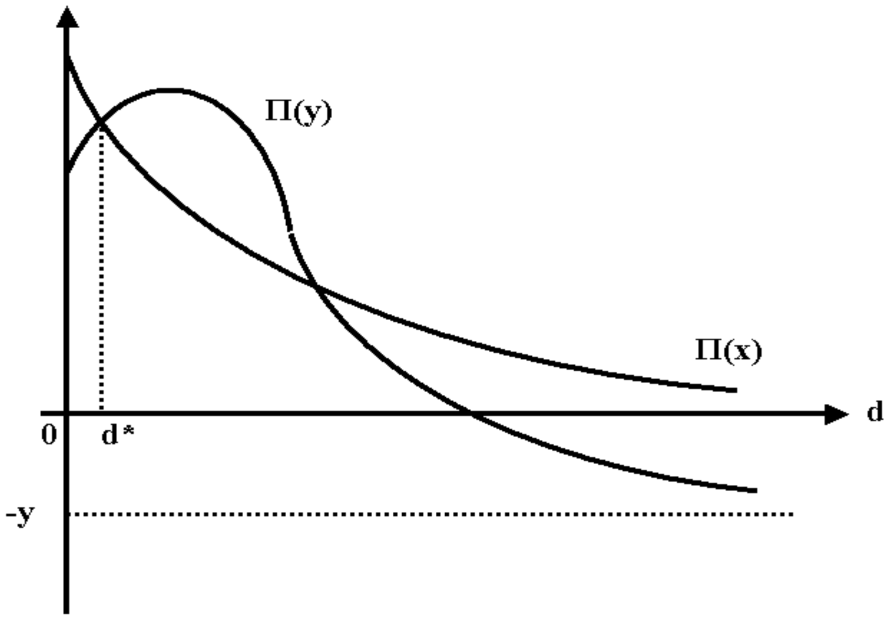

Figure 3.b.

a. Time

As the innovation cycle unfolds, uncertainty about the patentability of innovation decreases. To see how this affects the optimal delay d* we only need to determine how the Π(x) and Π(y) curves shown in figure 3.b. shift over time.

$$\frac{\partial \Pi(x)}{\partial t} = e^{-rd} \frac{\partial P}{\partial t} \pi(x) < 0$$

$$\frac{\partial \Pi(y)}{\partial t} = -e^{-rd} \frac{\partial P}{\partial t} \pi(y) > 0$$

The intuition for these results is straightforward: as scientific uncertainty decreases, the precision of the patent examiner 's decisions improves, which is good news for genuine innovations and bad news for illegitimate applications. Turning to figure 4, we see that, as time passes, Π(y) shifts upward and Π(x) downward. Hence the socially optimal review time, d*, is a decreasing function of t.

Proposition 2: The welfare maximising patent examination period is a decreasing function of t.

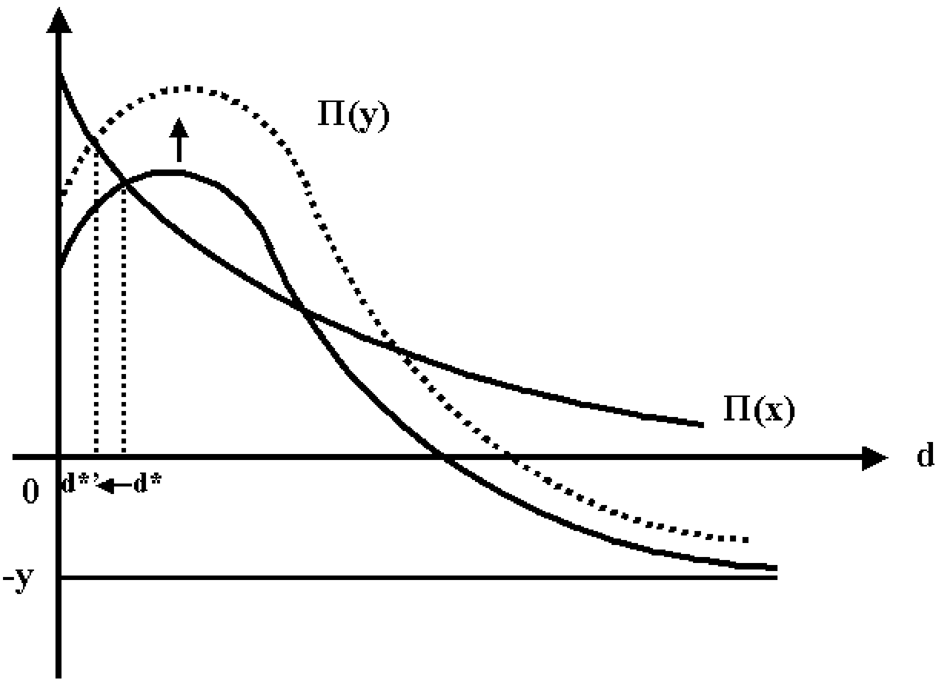

Figure 4

b. Importance of the Innovation

In our very stylised model, the most natural measure of patent "importance" is the value of the socially worthwhile innovations y. As there is no reason to expect that the proportion between social and private benefits would change as the importance of innovation changes, we simply redefine the *net* private value of innovation as $\alpha[\pi(y) - y]$ and the net social value of innovation as $\alpha[W(y) - y]$. The effect of an increase in α is then easily determined. Let us refer to figure 5. As α does not affect the private payoffs of an x investors, $\Pi(x)$ does not change. On the other hand, an increase in α 'squeezes $\Pi(y)$ closer to the horizontal axis, pushing it down for positive values of Π(y) and pulling it up for negative values of Π(y). The net result is, unambiguously, a decrease in d*.

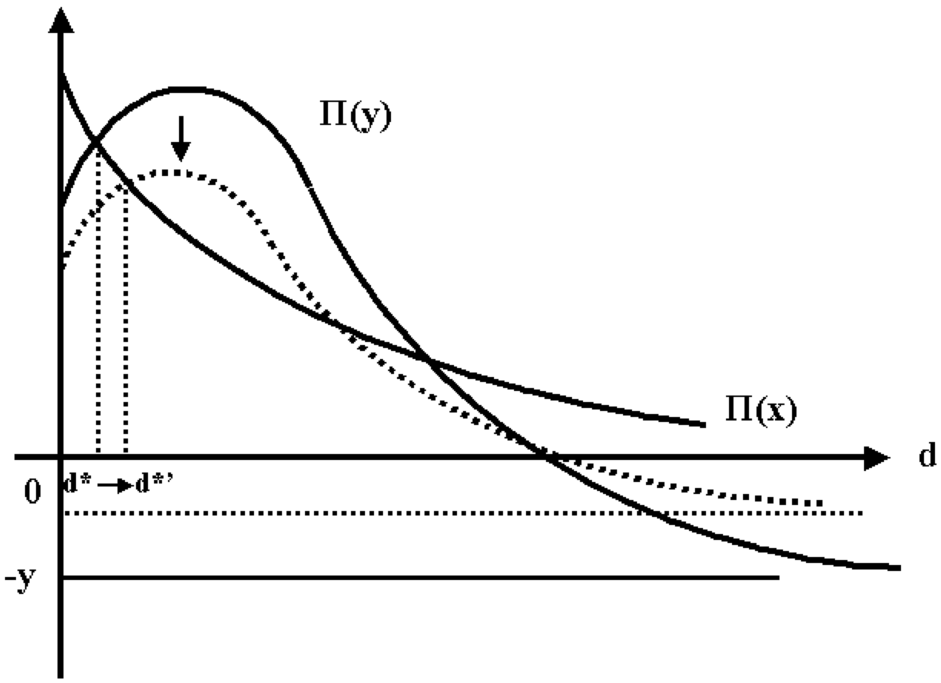

Figure 5

Proposition 3: The welfare maximising patent examination period is a decreasing function of the importance of the innovation.

III. Empirical Implications and Data

Combining propositions 1 through 3, yields some interesting conclusions. Firstly, with an optimal review scheme, "importance" per se should not be associated with longer delays. On the contrary more important patents should be approved more quickly. If on controls for the importance of patents, the length of the patent examination process should also *decrease* as the industry moves from the beginning of a new innovation cycle to a phase of lesser technological uncertainty. If one believes that more important innovations are more likely to emerge near the beginning of an innovation cycle, then the optimal review time can either increase or decrease over time. Formally, if we now specify $\alpha = \alpha(t)$, with $\alpha'(t) < 0$, the condition for d* to decrease over time is:

$$\alpha'(t)[-y + e^{-rd^*}(1-P)\pi(y)] - \alpha(t)e^{-rd^*}\frac{\partial P}{\partial t} > 0$$

Provided that this condition is satisfied, a cross-section analysis of the relationship between the importance of patents and the lag between their filing and grant dates could easily show a positive correlation, even though the lag is the result of an optimal policy. Moreover, the positive correlation would emerge, in spite of the fact that, other things equal, the socially optimal lag *decreases* as the importance of patents increases. To see this, consider two industries at different stages of their respective innovation cycles. At the time of the cross-section, industry A is further along in its cycle and is therefore characterised by both shorter delays and less important patents than industry B. Consequently, a cross section analysis would show that the length of the review process is positively related to the importance of the approved patent.

One way of avoiding such spurious correlation is to control for the fact that different observations belong to different ‘innovation cycles’. Unfortunately, this cannot be achieved by simply controlling for the patent class of the innovation – the most common approach – or, as in Johnson and Popp (2003), by grouping patents into five broad ‘technology groups’. Johnson and Popp ‘s control classification is far too coarse, while the conventional use of three digit patent classes is inappropriate. To appreciate this last point, an example might be helpful. Consider the recent wave of inventions in the field of genetically modified cops. The main patent classes attributed to the relevant inventions are 800, 435, 536 and, to a lesser extent, 530. The 800 class groups *all patents* concerning plants. This not only includes genetically engineered plants but also hybridisation techniques, and a host of other inventions useful for growing plants. On the other hand, the 435 class relates more closely to genetic manipulation but it embraces applications ranging from plants to animals, to human medicine. One should add that, as most patents are characterised by several patent classes, the choice of the primary class often appears to be somewhat arbitrary.

To minimise such pitfalls, we constructed a data set that only contains patents relating to a precisely defined field of innovation. The chosen field is that of genetically modified crops and we focus on patents granted by the USPTO. This has several advantages for our purpose. Firstly, although identifying the relevant patents is quite labour-intensive the criteria used to define the field are relatively simple: a patent must be about genetic engineering (as opposed to genetic manipulation through plant husbandry) of plant crops. These patents can be about the transformed plant itself, genes or other constructs used in the transformation process or genetic engineering techniques.[5] A second advantage is that GM crop innovation is a recent phenomenon with a readily identifiable beginning. The concept of 'early innovation' does not therefore present any difficulty.

In the current version of he paper, we will rely on patents granted between 1988 (the first US GM crop patent) and the end of 1998. This amounts to 635 observations. For each of these patents, we have both the filing date and the grant date and therefore the corresponding approval delay. We also need some measure of the patents ' "importance". One of the most commonly used measures, championed by Trajtenberg, Henderson and Jaffe (1997), is the total number of cites received by a given patent. In our data set, these cites were computed as of the beginning of 2003. This means that the latest patent in our sample could accumulate citations over a little more than four years. Alternatively, we will also use the number of claims in the patent document[6] and a measure of patent 'scope' based on a detailed reading of patent claims.

Finally the data set also include some useful control variables such as the identity of the examiner in charge of the patent, the identity of the attorney or agent who shepherded the application through the approval process and the type of the patent-holder (e.g. corporate versus university).

[5] Genetic engineering techniques are only considered if they claim a specific usefulness for the transformation of plants.

IV. Empirical Results: A Very First Look at the Data

We are interested in the relationship between the patent approval delay, EXATIME, the date at which the patent application was filed, FILE and a measure of the patent 's importance, IMP:

$$\text{EXATIME} = \alpha + \beta_1 \text{ FILE} + \beta_2 \text{ IMP} + u \qquad \text{..........................(1)}$$

If the examination delay results from an optimal patent review policy our model would predict that both β_1 and β_2 should be negative. The first column in Table 1 shows the OLS results for this simple-minded regression when IMP is measured as the total number of downstream cites. FILE is measured in months from December 1983. EXATIME is also measured in months. Robust standard error are included in bracket. Three stars indicate significance at the 0.01 level, two stars at the 0.05 level and one star at the 0.1 level.

As both FILE and CITES have strongly significant coefficient, the results appear to support the predictions of our model. They suggest that Johnson and Popp 's conclusion that more important patents experience longer delay is either the result of cross section effect described above or, at the very least, is not verified for the GM crop industry. As a benchmark, we also estimated the relationship between EXATIME and the number of citations received but without correcting for the position of the patents in the new innovation cycle. The results, presented in the second columns of Table 1, show that, in the absence of any time correction one would in fact conclude that more important patents experience longer delays. In other words, because important patents tend to emerge early, one would fail to distinguish between 'greater delay because of importance' and 'longer 'delay because it is early days'.

---

[6] . While there is little theoretical reason to believe that this ought to be a good measure of a patent 's

Table 1

CONSTANT	69.16*** (4,675)	32.4*** (0.625)	59.69*** (3,295)	61.03*** (3.211)	68.28*** (4,742)
FILE	-0.287*** (0.0335)		-0.232*** (0.0250)	-0.231*** (0.0249)	-0.273*** (0.0317)
CITES	-2.50** (0.1069)	0.109** (0.0559)			
CLAIMS			0.0707** (0.0333)		
SCOPE				-0.80 (0.946)	
IMPG					-0.228** (0.0995)
R^2	0.260	0.01	0.229	0.225	0.250

Unfortunately, the regression suffers from two major drawbacks. The first issue is that we do not observe the full distribution of citations received by the patents in our sample. Patent citations ' distributions typically have substantial right tails (see Caballero and Jaffe (1993)). In fact, even a casual look at the list of citing patents shows that even patents granted at the very beginning of the innovation cycle keep receiving significant numbers of new citations in 2002 and 2003. We must therefore accept the fact that CITE is a significantly censored measure of patent importance. Moreover, CITE will be more severely censored for later patents. In equation (1), β_2 measures the effect of CITE on the examination delay for a given file date. But for a given file date, a larger delay simply means a later grant date…. which means a more severely censored measure of the patent 's importance. In other words, the patent importance might well affect the examination delay but the examination delay also affects (negatively) our measure of patent importance. This would result in a negative estimate of β_2 even if patent importance had no effect on the approval lag. So, again,

importance, Lanjouw and Schankerman ( ) have shown that, in practice, it performs rather well.

we cannot safely assume that the negative estimate of β_2 obtained in Table 1 supports our predictions.

The second issue is that we only observe the patents that have been granted by the end of 1998. This might not be a concern for the earliest innovations, as we would expect most applications filed before 1990 to have been approved (or denied) by the end of 1998[7]. For later filing dates, however, the patent applications that find their way into our data set will precisely be those that have been approved relatively fast. This selection bias is stronger the later the filing date. This creates an artificial negative relationship between EXATIME and FILE, making it hazardous to claim the negative value for β_1 is supporting evidence for our model

We have tried to address the first issue in two ways. Firstly, and most simply, we have used alternative measures of patent importance that are not censored in our sample. The first alternative measure is the number of claims in the patent document. There is in fact no obvious reason for the number of claims to be a good indicator of a patent 's importance. After all, patent law specialists often insist that 'more is less', i.e. the more detailed the claims, the less general the invention. On the other hand, even if the 'more is less' principle applies to every single aspect of an invention, more significant innovations might have more aspects that are worth protecting and, therefore, more claims. Which of these two effects would prevail in practice is an empirical issue. Lanjouw & Schankerman (1999) show that the number of claims tends in fact to be a good indicator of the quality of patents. Our third measure of patent importance is a measure of patent 'breadth' made possible by the level of detail of our data set. As each potentially relevant patent was read, we were able to determine whether it claimed to modify a specific plant trait (e.g. herbicide resistance) and/or whether if claimed to apply to a particular plant or class of plants. Following the 'more is less' principle, less specific patents are classified as broader. We therefore created a SCOPE variable that takes the value 0 if the patent claims a specific trait or a specific plant and the value 1 otherwise.

[7] In our sample we only have 9 approval delays that are longer than 5 years and only 21 more that are longer than four years.

Turning back to Table 1 we see that if we rerun our regression with these two alternative measures of importance we see that the estimate of β_1 is dramatically affected: it is negative but insignificant when SCOPE is used and significantly positive when a patent 's importance is measured by its number of claims. There are however good reasons to believe that, in our sample, the number of claims is not a good measure of importance. Quite strikingly, the number of claims by patent increases rather sharply over time in spite of the fact that most of the fundamental innovations occurred relatively early on.[8]

We also tried to obtain a measure of CITE that is – roughly – corrected for censorship. This was done by estimating the effect of the file and grant dates on the total number of citations obtained[9] and then using this estimate to adjust the total number of cites for differences in the patents ' grant date. The resulting variable IMPG is positive if the patent has received more citations than would be expected given its grant date and is positive if the patent has under-performed. As shown in the last column of Table 1, IMPG has a significantly negative coefficient. Overall, then, we conclude that the evidence mildly supports our expectation that, once the patents ' positions in the innovation cycle is controlled for, the time of approval should actually decrease with the importance of the patent.

We now turn to the second difficulty. The best way of addressing the fact that we only observe patents that have already been granted would be to try to model the truncation process and then estimate the model with two-stage least squares. This is an approach that we have not yet pursued. Instead we have tried to get some feeling for the extent of the truncation problem. We first looked at all patents with filing dates before January 1, 1993. This means that, from the population of patents with such filedates, only patents with approval times of more than 6 years would not be part of our data set. Among these 156 patents, only 9 had experienced delays of more than 5 years and only 17 more had delays between 4 and 5 years. As 94 of the 156 patents actually have file dates before January 1991, this suggests that the truncation problem is not likely to be severe.

---

[8] There is in fact a negative correlation between the number of claims and the number of cites received!
[9] Using the negative binomial regression, we get CITE = 6.256*** - 0.00467 FILE – 0.0273*** GRANT.

We also re-estimated equation (1) on two sub-samples, one where we only consider patents with the 376 patents with file dates before January 1995 and another where we only consider the 243 patents with file dates before January 1994. The results are shown in the second and third columns of tables 2.a., 2.b. and 2.c. The first column just restates the results of the full sample regression to facilitate comparisons. The three tables differ only in the measures of patent importance used.

Table 2.a.

	Full Sample	**File < Jan 94**	**File < Jan 93**	**File < Jan 94 & Exatime < 4 years**	**File < Jan 93 & Exatime < 5 years**
CONSTANT	69.16*** (4,675)	68.06*** (5.633)	68.79*** (6.535)	44.45*** (2.736)	52.37*** (3.507)
FILE	-0.287*** (0.0335)	-0.285*** (0.0447)	-0.302*** (0.0577)	-0.112*** (0.0224)	-0.163*** (0.0326)
CITE	-2.50** (0.1069)	-0.224** (0.104)	-0.197* (0.103)	-0.0993* (0.0582)	-0.157** (0.064)
R^2	0.260	0.198	0.174	0.066	0.098

Comparing columns two and three to the first shows that the results remain remarkably similar. In particular, the coefficient of FILE remains significantly negative at the 0.01 level and does not decrease in magnitude when the sample is limited to patents with early file dates. This again suggests that the possible bias generated by the truncation of our data might not be too large.

Table 2. b.

	Full Sample	File < Jan 94	File < Jan 93	File < Jan 94 & Exatime < 4 years	File < Jan 93 & Exatime < 5 years
CONSTANT	61.03*** (3.211)	61.16*** (4.051)	62.88*** (4.763)	41.7*** (2.124)	48.11*** (2.841)
FILE	-0.231*** (0.0249)	-0.234*** (0.0349)	-0.254*** (0.0465)	-0.0902*** (0.0186)	-0.129*** (0.028)
SCOPE	-0.80 (0.946)	-1.37 (1.454)	-2.00 (1.968)	-1.444 (1.00)	-1.875 (1.494)
R^2	0.225	0.168	0.149	0.06	0.079

Table 2.c.

	Full Sample	File < Jan 94	File < Jan 93	File < Jan 94 & Exatime < 4 years	File < Jan 93 & Exatime < 5 years
CONSTANT	68.28*** (4,742)	67.18*** (5.713)	67.97*** (6.627)	43.8*** (2.77)	51.77*** (3.583)
FILE	-0.273*** (0.0317)	-0.272*** (0.0432)	-0.29*** (0.0563)	-0.105*** (0.0216)	-0.154*** (0.032)
IMPG	-0.228** (0.0995)	-0.19* (0.0976)	-0.167* (0.976)	-0.079 (0.0573)	-0.134** (0.0632)
R^2	0.250	0.189	0.166	0.062	0.091

Finally we also estimated equation (1) on the two sub-samples just described, where we further restricted our attention to patents that would have been approved within the maximum approval 'window' available to the latest patent in the subsample. For example, the data set with file dates before January 1995 was further restricted to patents that were indeed approved within four years. This ensures that patents effectively experience the same truncation irrespective of their actual filing date. The results are shown in the last two columns of tables 2.a., 2.b. and 2.c. This time, the magnitude of the FILE coefficient is reduced – by about half- but it remains significantly negative suggesting that the examination delay truly tends to be larger for patents that are filed earlier in the innovation cycle. In fact the magnitude of the FILE coefficient remains non-negligible. Even if we take the lowest estimates obtained, patents filed one year later will, on average, experience approval delays that are shorter by one month.

We also hope to eventually address the dependent variable truncation issue by simply extending the data set to include all patents granted until the end of 2002. This would allow us to consider only patents granted before 1996, so that only patents experiencing delays of more than seven years would be omitted, and still have a sufficiently large sample.

A final difficulty comes from the fact that the patent review process is not in fact bound by very precise rules about either the length of the examination period or the types of claims that should be approved. What we call patent 'policy' is the result of a large number of interactions between a relatively small number of patent examiners and a variety of patent applicants and their representatives. It would therefore be comforting to know that our results are robust to the possible idiosyncratic behaviour of a few examiners and/or applicants. We have therefore augmented our basic regression with three groups of control variables. A total of ten dummies were defined to account for the behaviour of patent each patent examiner who reviewed at least 10 patents of our sample. Similarly, nineteen dummies were assigned to attorneys who had helped with the filing of at least ten (successful) patent applications. Finally we also classify the patent holders into five groups: corporations, individuals, government, universities and 'Mixed', including dummies for the last four. The 'mixed' category includes corporations located near a university with major GM crop research activities (e.g. Calgene and UC Davis).

The results are shown in Table 3. Given the number of dummies, we do not show the coefficient on the individual examiner and attorney dummies, even though some of these are significant. What matters for our purpose is that each set of dummy is jointly significant in each of the regressions presented. As always, numbers in bracket are robust standard errors. Table 4 shows the same regressions when we only consider patents filed before 1995 and Table 5 shows the same regressions when we only select patents with a time of approval no greater than four years.

Table 3

CONSTANT	72.42*** (5.538)	62.82*** (4.185)	64.87*** (4.13)	71.56*** (5.61)
FILE	-0.343*** (0.103)	-0.294*** (0.0315)	-0.292*** (0.0314)	-0.33*** (0.038)
CITE	-0.251** (0.103)			
CLAIMS		0.090*** (0.0344)		
SCOPE			-1.18 (0.965)	
IMPG				-0.214** (0.0954)
UNIV	2.36* (1.246)	2.916** (1.262)	2.69** (1.261)	2.41* (1.248)
INDIV	2.73 (2.694)	4.21 (2.834)	3.84 (2.834)	2.94 (2.717)
GOVT	3.51* (1.898)	4.32** (1.941)	4.16** (2.017)	3.64* (1.916)
MIXED	2.23 (1.534)	3.15** (1.540)	2.77* (1.544)	2.32 (1.542)
EXA	Signif.	Signif.	Signif	Signif.
ATTORNEY	Signif.	Signif.	Signif	Signif.
R^2	0.355	0.333	0.331	0.347

Table 4

CONSTANT	71.97*** (6.927)		65.42*** (5.390)	71.22*** (7.0267)
FILE	-0.354*** (0.584)	-0.305*** (0.0502)	-0.306*** (0.0498)	-0.34*** (0.057)
CITE	-0.227** (0.1048)			
CLAIMS		0.112** (0.0547)		
SCOPE			-2.02 (1.498)	
IMPG				-0.196** (0.0984)
UNIV	2.481 (1.844)	3.205* (1.864)	3.02 (1.871)	2.54 (1.847)
INDIV	1.806 (4.019)	3.45 (4.281)	2.84 (4.298)	2.02 (4.056)
GOVT	2.764 (2.501)	3.50 (2.677)	3.41 (2.729)	2.90 (2.528)
MIXED	0.316 (2.324)	1.15 (2.334)	0.75 (2.314)	0.38 (2.34)
EXA	Signif.	Signif.	Signif.	Signif.
ATTORNEY	Signif.	Signif.	Signif.	Signif.
R^2	0.337	0.314	0.311	0.329

Table 5

CONSTANT	40.38*** (3.176)	37.07*** (2.724)	39.46*** (2.811)	39.73*** (3.169)
FILE	-0.121*** (0.0253)	-0.112*** (0.0226)	-0.111*** (1.06)	-0.116*** (0.0244)
CITE	-0.062 (0.061)			
CLAIMS		0.0882** (0.0373)		
SCOPE			-1.65 (1.061)	
IMPG				-0.044 (0.061)
UNIV	3.057** (1.331)	3.51*** (1.336)	3.34** (1.342)	3.11** (1.336)
INDIV	2.66 (3.034)	2.93 (3.11)	2.31 (3.048)	2.72 (3.044)
GOVT	4.91** (2.308)	5.02** (2.284)	4.82** (2.327)	4.98** (2.314)
MIXED	2.38 (1.529)	2.83* (1.546)	2.59* (1.502)	2.44 (1.528)
EXA	Signif.	Signif.	Signif.	Signif.
ATTORNEY	Signif.	Signif.	Signif.	Signif.
R^2	0.244	0.253	0.246	0.242

Comparing tables 3,4 and 4 to table 1 and 2, shows that the estimates of β_1 and β_2 do not change much. They remain significantly negative whenever they were negative before. If anything the magnitude of the estimated coefficients increases slightly. The estimated coefficients of the UNIV, INDIV, GOVT and MIXED dummies are of some independent interest. There is some evidence that government agencies and – especially – universities get their patents through the examination process more slowly than private corporations. One can think of at least three reasons for this. Firstly, universities might work on different sub-fields where, for technical reasons, approval takes longer. Secondly, universities might just have less experience than corporation in pushing their applications through the patent office. Finally, because of differences in internal incentives, universities might just not work as hard at getting their applications approved quickly. While of full discussion of this issue goes beyond the scope of this paper, there are good reasons to believe that the 'incentive' explanation is the most convincing. To evaluate the first explanation, we ran similar regressions with additional dummies controlling for the specific traits and plants involved in the patents, trying to get at the notion that universities might be working on 'different things'. This did not significantly affect the coefficient of the university dummy. The learning-based explanation is a priori appealing but a large proportion of the university patents in our sample are held by large research universities that have as much experience with the USPTO than most large corporations.

V. Conclusion

Innovations are often the subject of intense scrutiny. In the U.S., for example, drugs must be reviewed by the FDA before being approved for prescription or over the counter sale. Inventors seeking to protect their intellectual property by patents must also go through a detailed review by the Patent Office. The main purpose of this review is to determine whether the alleged inventor has actually produced something novel, non-obvious and useful. These administrative processes can take significant time. This raises the question of whether – by design or inadvertently – they treat different types of innovations differently and whether such discrimination is desirable. Of particular concern is whether more important innovations tend to be approved more slowly. In the cases of prescription drugs, further delays for important medicines would have a large social loss in terms of missed opportunities for more effective treatment. Dranove and Metzler (1994) argue that concerns about increased delays of approval at the US FDA might be exaggerated as important drugs do get approved significantly faster than others. On the other hand, Johnson & Popp find that more significant innovations tend to be approved more slowly by the US patent office. Under the new legal regime, patent applications, and therefore the information required to understand the invention, must be published after eighteen months. The greater delay experienced by more significant patents could therefore undermine private incentives to pursue major innovations. Unless the social benefits from accelerated diffusion of knowledge are large, the combination of early disclosure and bias in the approval time will reduce welfare.

The first part of our paper develops a simple model of the patent review process that emphasises another role of the examination delay. In our model, longer examination time makes for more accurate decisions: it reduces the proportion of legitimate inventions that are rejected and increases the proportion of illegitimate inventions that are mistakenly approved. This, in turn, increases the inventors ' incentives to undertake legitimate research. Moreover there is learning in the sense that the precision of the examiners ' decisions also increase as the general technology behind a given innovation wave becomes more familiar. We solve this model for the welfare-maximising time of patent application review and obtain two main predictions. Firstly, controlling for the importance of useful innovations, the optimal delay decreases over time as technological uncertainty is reduced. Secondly, controlling for the level of technological uncertainty, the delay should decrease when the importance of useful innovations increases. While our model is very simple and specific, its main results are quite robust. They would in fact obtain in almost any model where 'important' research is costlier but socially more desirable than 'unimportant' research and where longer delays means more accurate decisions.

Our finding also has an interesting implication for empirical work. If one believes that more important innovations tend to occur at the beginning of a new technological wave then using patent data from a broad cross- section of industries can easily result in a spurious positive correlation between approval delay and patent importance. In other words, a positive correlation could be obtained even though the patent office, behaving according to our optimal policy, would, other things equal, approve important patents more quickly.

In the second part of the paper, we use a data set on US GM crop patents to determine whether the timing of patent approval is consistent with our predictions.

The main advantage of the data set is that, as we capture a well - defined technological 'wave' from its beginning, we do not have to worry about any complication coming from the cross-sectional nature of the data. Although the empirical analysis presented is still rather rough and the results must therefore be taken as preliminary, there seems to be significant evidence in support of our predictions. In particular, the decrease in approval times as the new technological wave unfolds is not only significant but also rather large: depending on the precise regression considered, patents applications filed one year later are approved between one and three months faster on average. The larger number amounts to about 8% of the median approval lag in the sample. The evidence on the importance of patents is weaker. We generally obtain the expected negative relationship when we measure importance as the number of downward cites but the number of claims in a patent results in a positive or an insignificant coefficient. Using an alternative measure of patent scope consistently yields a negative coefficient but it is not significant at the 0.1 level.

It should be emphasised that our empirical results should, at this stage, been considered as a first look at the data. Not only would we like to revisit the results on a slightly longer data set but it would also be useful to estimate the determinants of the patent approval delay as a hazard rate model. Early results suggest that this later extension only strengthens our conclusions.

References

Caballero, R.J. and A.B. Jaffe, 1993, "How High Are the |Giants ' Shoulders: An Empirical Assessment of Knowledge Spillovers and Creative destruction in a Model of Economic Growth", in O.J. Blanchard and S. Fischer, eds, *NBER Macroeconomics Annual*. MIT Press, Cambridge, Mass.

Dranove, D. and D. Metzler, 1994, "Do Important Drugs Reach the Market Sooner?", *Rand Journal of Economics*, Autumn, pp. 402-423.

Lanjouw, J.O. Nad M. Schankerman, 1999, "The Quality of Ideas: Measuring Innovation with Multiple Indicators", *NBER Working Paper No 7345*.

Johnson, D.K.N. and D. Popp, 2003, "Forced Out of the Closet: The Impact of the American Inventors Protection Act on the Timing of Patent Disclosure", *Rand Journal of Economics*, Spring, Vol 34.1, pp. 96 – 112.

Appendix

Conditions for Π(y) to have a single maximand d_{max}.

Given the assumptions made in the text we only need to further assume that $\frac{\partial^3 P}{\partial d^3} < 0$ and $\lim_{d\to\infty}.\frac{\partial^2 P}{\partial d^2} = 0$.

We already know that $\frac{\partial \Pi(y)}{\partial d}(d = 0) > 0$ and that Π(y) approaches –y asymptotically from above as d goes to infinity. Moreover, Π(y) is continuous in d. Hence, if we can show that Π(y) is concave at d = 0 and changes convexity at most once, then we will have established that it has a single interior maximum.

We have

$$\frac{\partial \Pi(y)}{\partial d} = \pi(y)e^{-rd}[-r(1-P(d,t)) - \frac{\partial P}{\partial d}]$$

$$\frac{\partial^2 \Pi(y)}{\partial d^2} = r^2(1-P) + 2r\frac{\partial P}{\partial d} - \frac{\partial^2 P}{\partial d^2} \qquad \text{[A1]}$$

Since $\frac{\partial^2 P}{\partial d} > 0$ and $-\frac{\partial P}{\partial d}(d = 0) > \frac{r}{2}$

$$\frac{\partial^2 \Pi(y)}{\partial d^2}(d = 0) = r^2(1-P) + 2r\frac{\partial P}{\partial d} - \frac{\partial^2 P}{\partial d^2} < 0$$

$$\lim_{d\to\infty}\frac{\partial^2 \Pi(y)}{\partial d^2} = r^2 + 2r\lim_{d\to\infty}\frac{\partial P}{\partial d} - \lim_{d\to\infty}\frac{\partial^2 P}{\partial d^2} = r^2 > 0$$

Moreover, as d increases, the first (positive)term in equation A1, increases as P decreases. On the other hand, under our assumptions, the absolute value of the two other (negative) terms decreases. Hence Π(y) is initially concave in d and changes convexity only once.